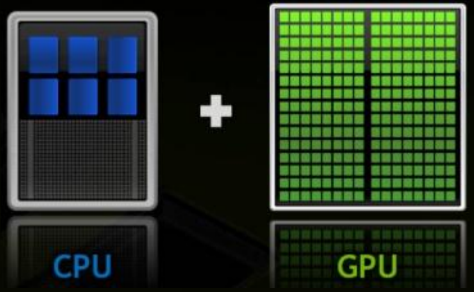
Status and Directions of NVIDIA GPUs for Earth System Modeling

Stan Posey | HPC Industry Development

NVIDIA, Santa Clara, CA, USA

NVIDIA and HPC Evolution of GPUs

- Public, based in Santa Clara, CA | ~$4B revenue | ~5,500 employees
- Founded in 1999 with primary business in semiconductor industry
  - Products for graphics in workstations, notebooks, mobile devices, etc.
  - Began R&D of GPUs for HPC in 2004, released first Tesla and CUDA in 2007
- Development of GPUs as a co-processing accelerator for x86 CPUs



HPC Evolution of GPUs

- **2004:** Began strategic investments in GPU as HPC co-processor
- **2006:** G80 first GPU with built-in compute features, 128 cores; CUDA SDK Beta
- **2007:** Tesla 8-series based on G80, 128 cores – CUDA 1.0, 1.1
- **2008:** Tesla 10-series based on GT 200, 240 cores – CUDA 2.0, 2.3
- **2009:** Tesla 20-series, code named “Fermi” up to 512 cores – CUDA SDK 3.0, 3.2

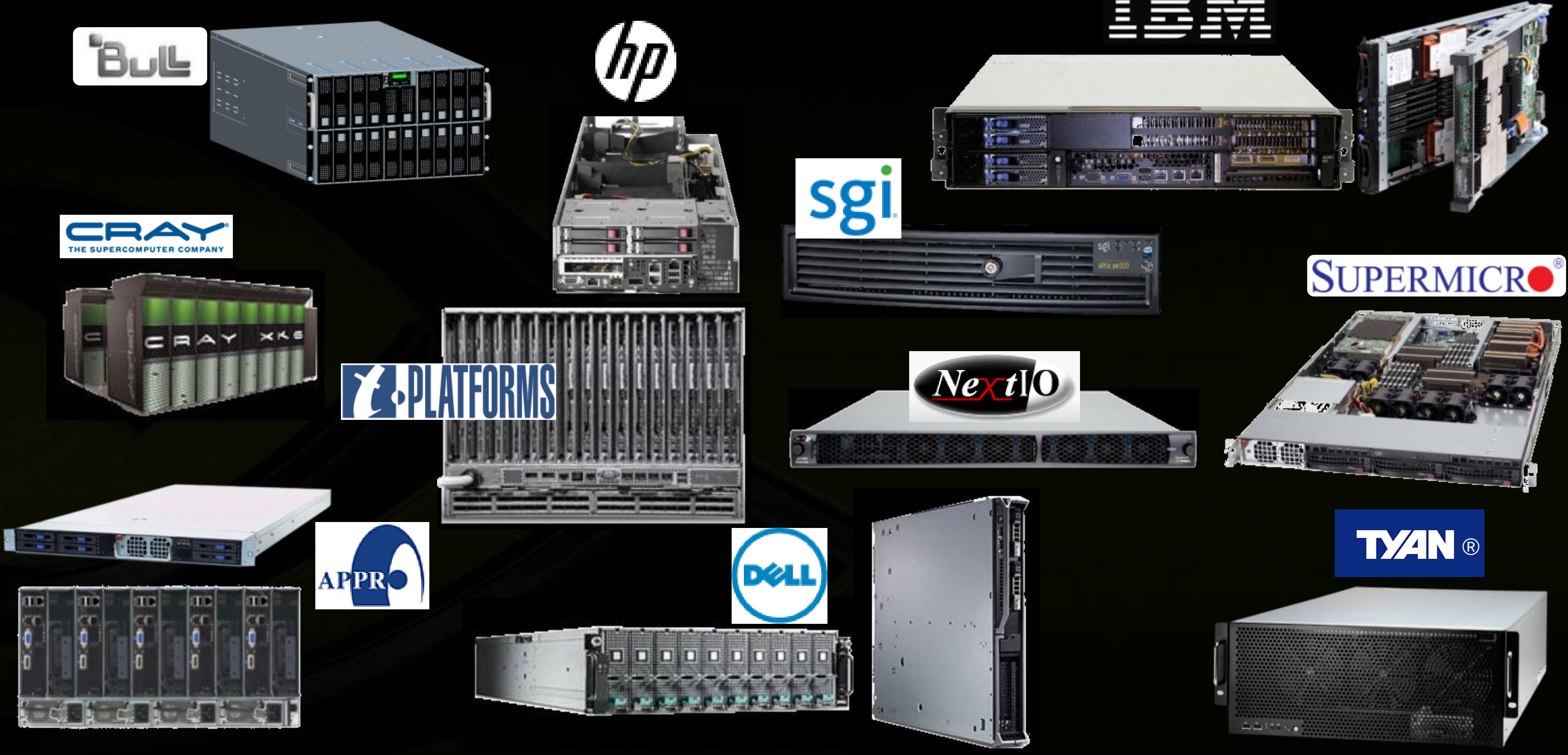
3 Years With
3 Generations

Agenda: Status and Directions of NVIDIA GPUs

- GPU Developments for Earth System Modeling
- Application Development Considerations
- GPU Hardware and Software Roadmap

GPUs Are Now Mainstream HPC

System Vendors Who Provide Tesla M-series GPUs



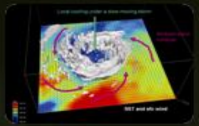
GPUs Are Now Mainstream HPC

NVIDIA

Oil & Gas

- Schlumberger
- BR Petrobras
- Total
- Chevron
- Paradigm

Edu/Research

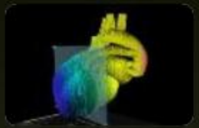


- Chinese Academy of Sciences
- Georgia Tech
- Harvard School of Engineering and Applied Sciences
- Oak Ridge National Laboratory

Government

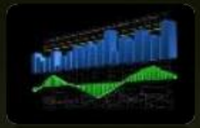
- Air Force Research Laboratory
- NASA
- Naval Research Laboratory
- BAE Systems

Life Sciences



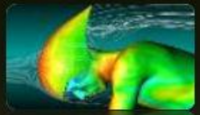
- Boston Scientific
- MGH 1811 Mass General Hospital
- Max Planck Institute
- Beckman Coulter

Finance



- Bloomberg
- BNP Paribas
- Standard Life
- SciComp Inc.
- NumeriX

Manufacturing



- Agilent
- ANSYS
- Autodesk
- SIMULIA
- ACUSIM Software

insideHPC **Nvidia CEO Jen-Hsun Huang to Keynote SC11**

NVIDIA GPUs Power 3 of Top 5 Supercomputers

#2 : Tianhe-1A

7168 Tesla GPU’s 2.5 PFLOPS

#4 : Nebulae

4650 Tesla GPU’s 1.2 PFLOPS

#5 : Tsubame 2.0

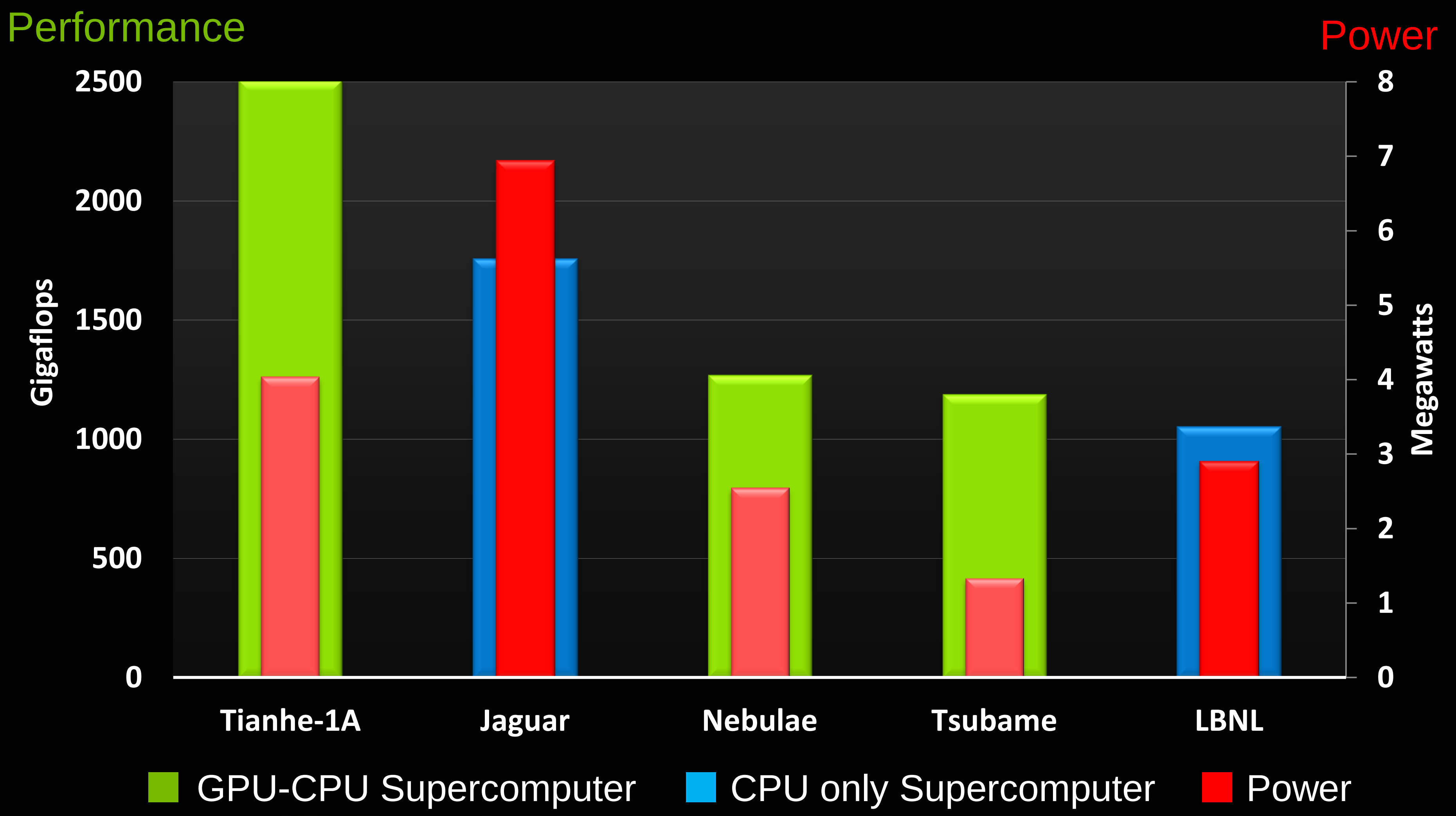
4224 Tesla GPUs 1.194 PFLOPS

“We not only created the world's fastest computer, but also implemented a heterogeneous computing architecture incorporating CPU and GPU, this is a new innovation.”

Premier Wen Jiabao

Public comments acknowledging Tianhe-1A

GPU Systems: More Power Efficient (for HPL)



Comparison with Top Supercomputer *K* in Japan

K Computer: Custom SPARC Processors

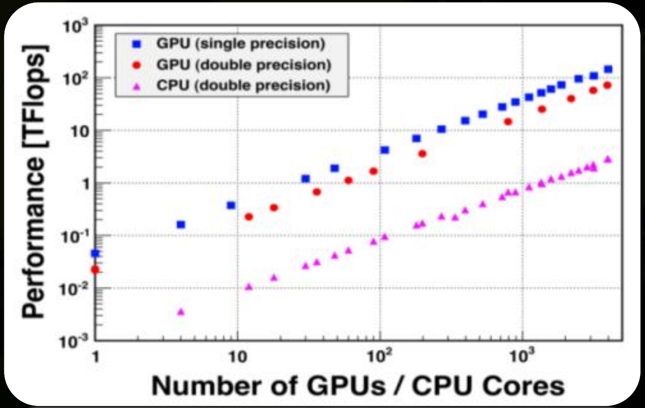
Tsubame: Intel CPUs + NVIDIA Tesla

K Computer		Tsubame
8.1 PetaFlop		1.2 PetaFlop
68,500 CPUs		2K CPUs, 4K GPUs
672 Racks	2.3x better flops/rack	44 Racks
10 Megawatt	1.06x better flop/watt	1.4 Megawatt
$700 Million	2.6x better $/flop	$40 Million

Science on GPUs: ASUCA NWP on Tsubame 2.0

Tsubame 2.0
Tokyo Institute of Technology

- 1.19 Petaflops
- 4,224 Tesla M2050 GPUs



3990 Tesla M2050s
145.0 Tflops SP
76.1 Tflops DP

Simulation on Tsubame 2.0, TiTech Supercomputer

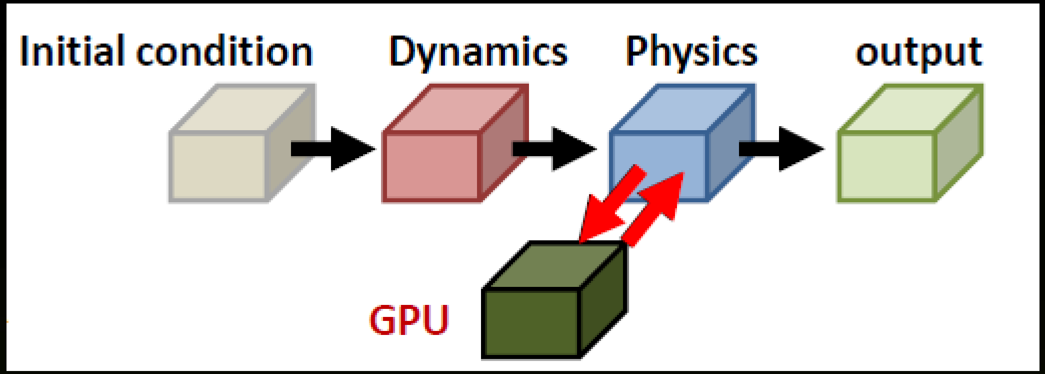
GPU Performance Requires Full GPU Approach

Physics Only

- WRF

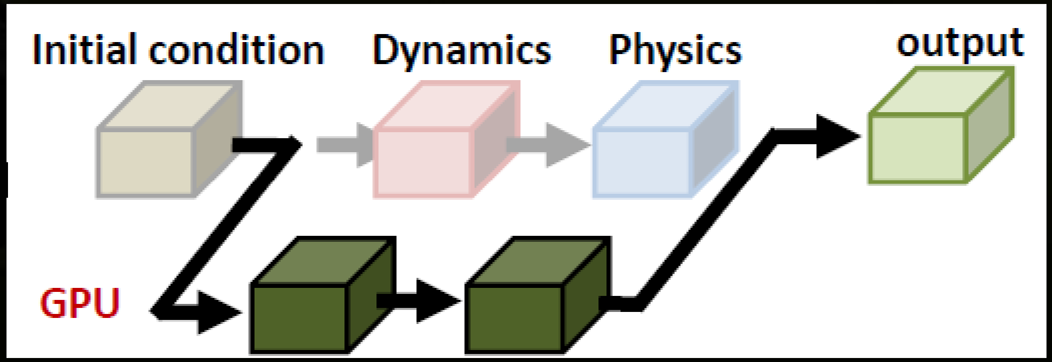
Dynamics Only

- HIRLAM
- HOMME



Full GPU Approach

- ASUCA
- CAM5
- COSMO
- GEOS-5
- GRAPE
- ICON
- NIM



NVIDIA Features GPU Progress at Conferences

Supercomputing 2010 | Nov 2010 | New Orleans, LA

COSMO: GPU Considerations for Next Generation Weather Simulations
Thomas Schulthess, Swiss National Supercomputing Centre (CSCS)

ASUCA: Full GPU Implementation of Weather Prediction Code on TSUBAME Supercomputer
Takayuki Aoki, GSIC of Tokyo Institute of Technology (TiTech)

NIM: Using GPUs to Run Next-Generation Weather Models
Mark Govett, National Oceanic and Atmospheric Administration (NOAA)

***BoF*:** GPUs and Numerical Weather Prediction *(organized by CSCS and NVIDIA)*
Featured organizations: TiTech (ASUCA), NASA (GEOS-5), NOAA (NIM), Cray, PGI

NVIDIA GPU Technology Conference | Sep 2010 | San Jose, CA

ASUCA: Full GPU Implementation of Weather Prediction Code on TSUBAME Supercomputer
Takayuki Aoki, GSIC of Tokyo Institute of Technology (TiTech)

NIM: Using GPUs to Run Next-Generation Weather Models
Mark Govett, National Oceanic and Atmospheric Administration (NOAA)

MITgcm: Designing a Geoscience Accelerator Library Accessible from High Level Languages
Chris Hill, Massachusetts Institute of Technology (MIT)

GPU Progress Reported at This Workshop

NVIDIA

Programming weather, climate, and earth-system models on heterogeneous multi-core platforms

September 7-8, 2011 at the National Center for Atmospheric Research in Boulder, Colorado

GPU related talks (11+) that cover application software such as:

NIM | WRF | GEOS-5 | HOMME | COSMO | CAM | ICON

Successes and Challenges using GPUs for Weather and Climate Models	Mark Govett, NOAA
Experience using FORTRAN GPU Compilers with the NIM	Tom Henderson, NOAA
GPU Acceleration of the RRTM in WRF using CUDA FORTRAN	Greg Ruetsch, NVIDIA
Lessons Learned adapting GEOS-5 GCM Physics to CUDA FORTRAN	Matt Thompson, NASA
Accelerated Cloud Resolving Model in Hybrid CPU-GPU Clusters	Jose Garcia, NCAR
Reworking Boundary Exchanges in HOMME for Many-Core Nodes	Ilene Carpenter, NREL
Performance optimizations for running an NWP model on GPUs	Jacques Middlecoff, NOAA
Rewrite of the COSMO Dynamical Core	Mueller / Gysi, SCS/CSCS
Experiences with the Finite-Volume Dynamical core and GEOS-5 on GPUs	Bill Putman, NASA
Progress in Accelerating CAM-SE	Jeff Larkin, Cray/ORNL
Porting the ICON Non-hydrostatic Dynamical Solver to GPUs	Will Sawyer, CSCS

Agenda: Status and Directions of NVIDIA GPUs

- GPU Developments for Earth System Modeling
- **Application Development Considerations**
- GPU Hardware and Software Roadmap

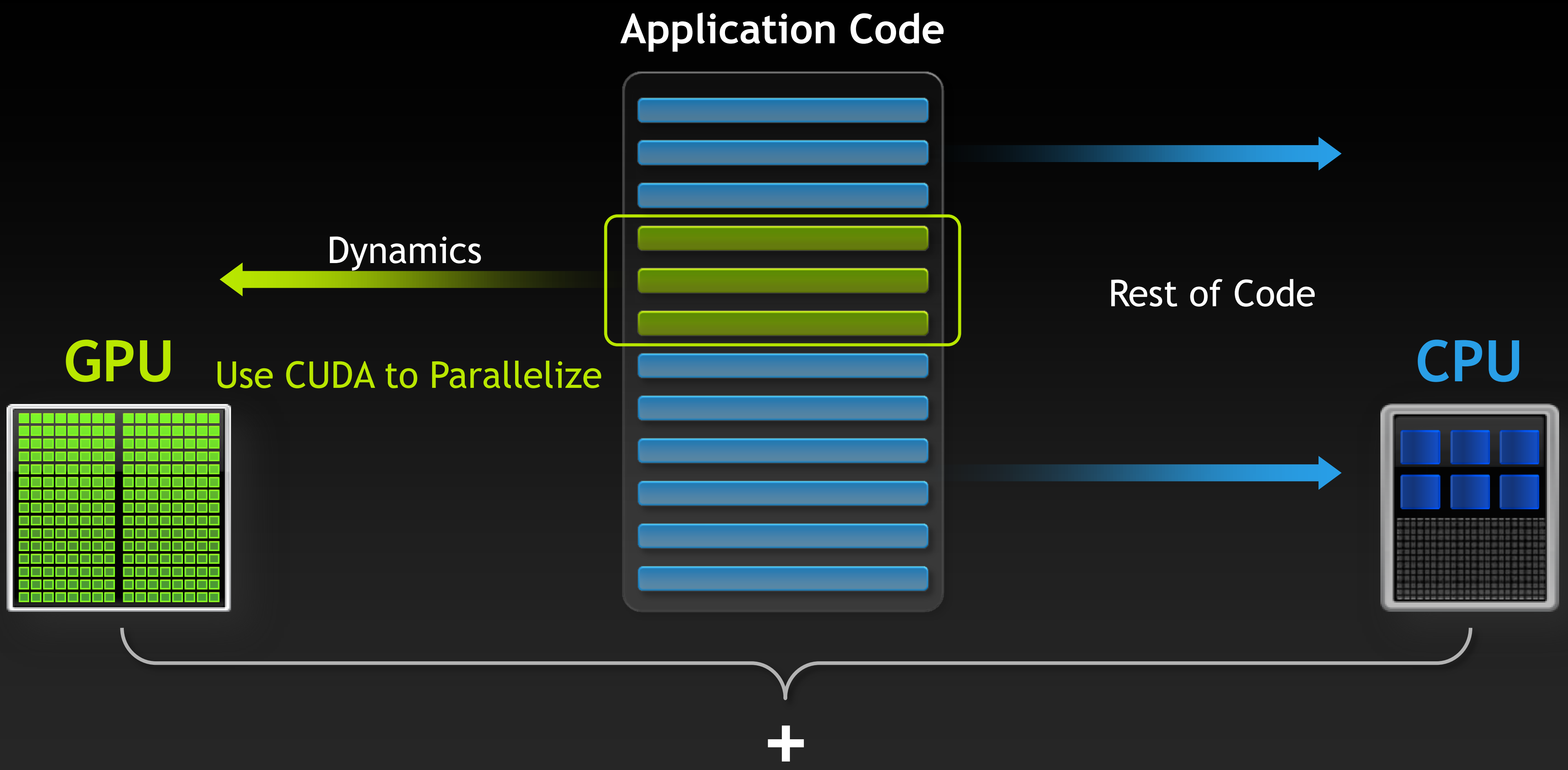
GPU Considerations for Climate-Weather Models

- Initial efforts are mostly dynamical core developments
  - If dynamics ~50% of profile time – 2x overall speed-up is possible
  - More of application must be moved to GPUs for full benefit
  - *Implicit schemes* – iterative sparse matrix linear algebra solvers
  - *Explicit schemes* – no linear algebra, operations on i,j,k stencil
- Increasing use of high performance GPU-based libraries
  - Examples: SpMV for iterative solvers; FFT for spectral methods
- Most models are parallel and scale across multiple CPU cores
  - Multi-core CPUs contribute to parallel matrix assembly, others
- Most models use a domain decomposition parallel method
  - Fits GPU model very well and preserves costly MPI investment
- Fortran programming on GPUs most critical for adoption
  - NVIDIA alliances and investments in CAPS, PGI and Cray compilers

Options for Parallel Programming of NVIDIA GPUs

Approach	Examples
Applications	MATLAB, Mathematica, LabVIEW
Libraries	FFT, BLAS, SPARSE, RNG, IMSL, CUSP, etc.
Directives	PGI Accelerator, HMPP, Cray
Wrappers	PyCUDA, CUDA.NET, jCUDA
Languages	CUDA C, CUDA C++, PGI CUDA Fortran, GPU.net
APIs	CUDA, OpenCL

Most Implementations Focus on Dynamical Core



Developments in Iterative Solvers for Implicit

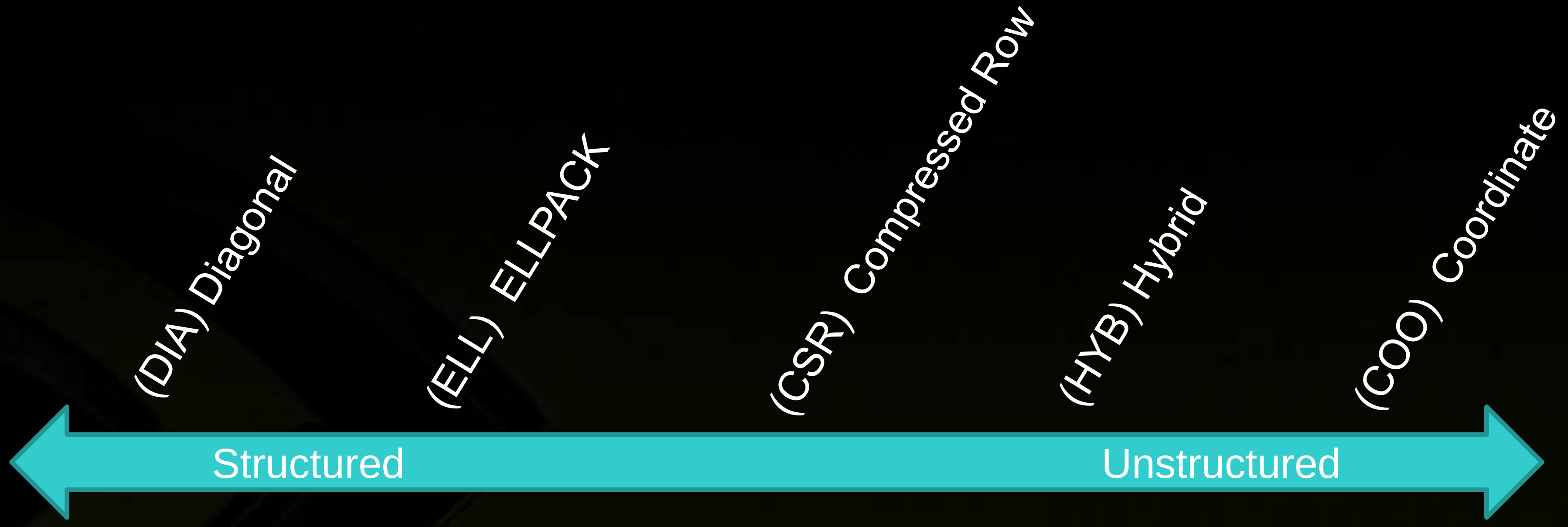
- **Sparse-matrix vector multiply (SpMV) & BLAS1**
  - **Memory-bound**

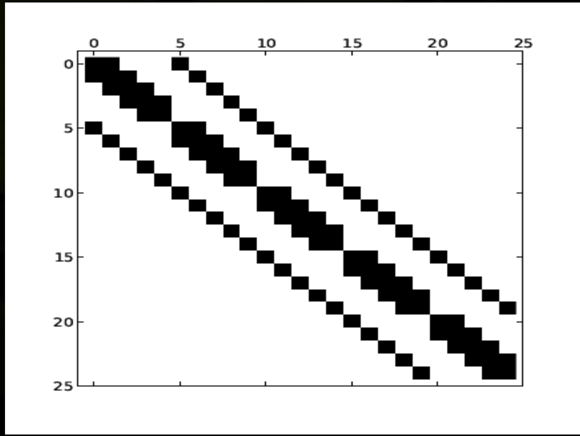
- **GPU can deliver good SpMV performance**
  - **~10-20 Gflops for unstructured matrices in double precision**

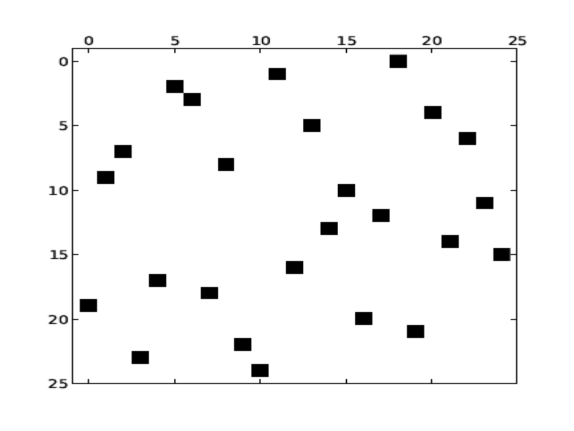
- **Best sparse matrix data structure on GPU different from CPU**
  - **Explore for your specific case**

- **A massively parallel preconditioner is key:**
  - Lectures: **Jon Cohen at IMA Workshop:** [“Thinking parallel: sparse iterative solvers with CUDA”]
  - **Nathan Bell (4-parts) at PASI:** [“Iterative methods for sparse linear systems on GPU”]

Typical Sparse Matrix Formats

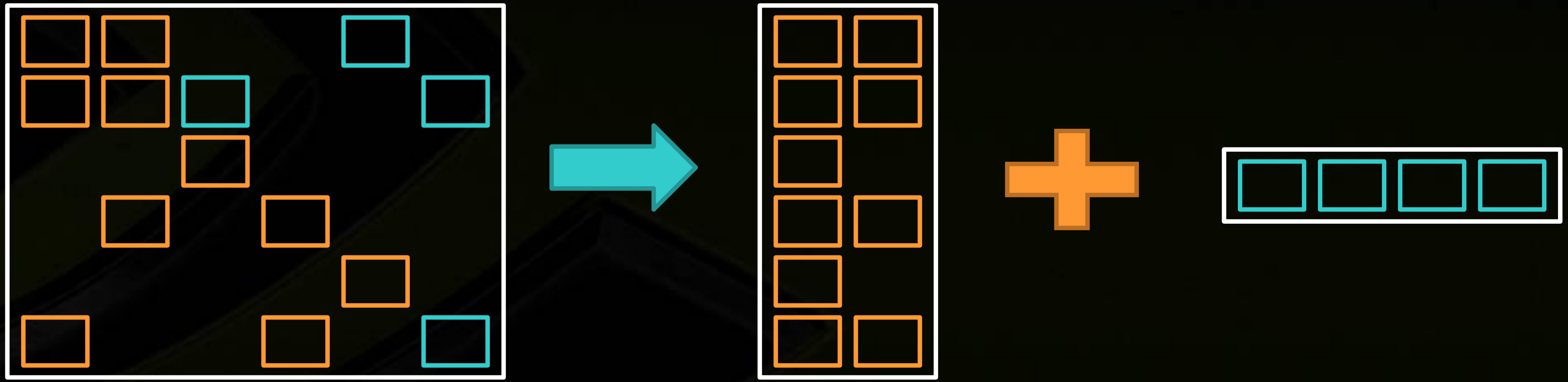






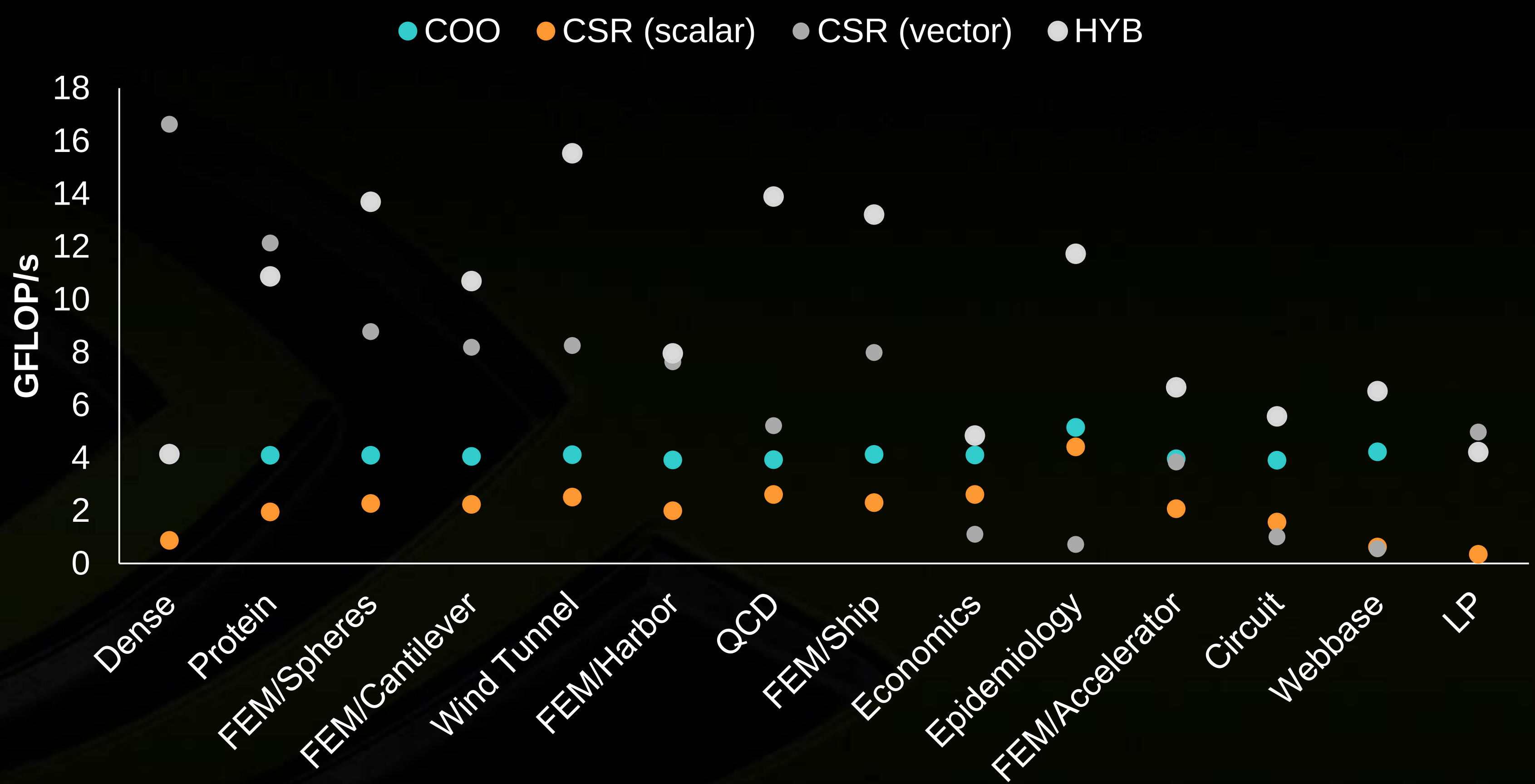
Hybrid Sparse Matrix Format for GPUs

- ELL handles *typical* entries
- COO handles *exceptional* entries
  - Implemented with segmented reduction



- Some overheads in matrix format conversion, can be hidden if the solver has O(100) of iterations

SpMV Performance for Unstructured Matrices



Flops=2*nz/t, BW = (2*sizeof(double)+size(int))/t

=> Flops~BW/10~15 Gflop/s

Agenda: Status and Directions of NVIDIA GPUs

- GPU Developments for Earth System Modeling
- Application Development Considerations
- **GPU Hardware and Software Roadmap**

NVIDIA Announced “Project Denver” Jan 2011

NVIDIA Announces "Project Denver" to Build Custom CPU Cores Based on ARM Architecture, Targeting Personal Computers to Supercomputers

NVIDIA Licenses ARM Architecture to Build Next-Generation Processors That Add a CPU to the GPU

LAS VEGAS, NV -- (Marketwire) -- 01/05/2011 -- **CES 2011** -- NVIDIA announced today that it plans to build high-performance ARM® based CPU cores, designed to support future products ranging from personal computers and servers to workstations and supercomputers.

It's true folks, NVIDIA's building a CPU! Madness! The future just got a lot more exciting.

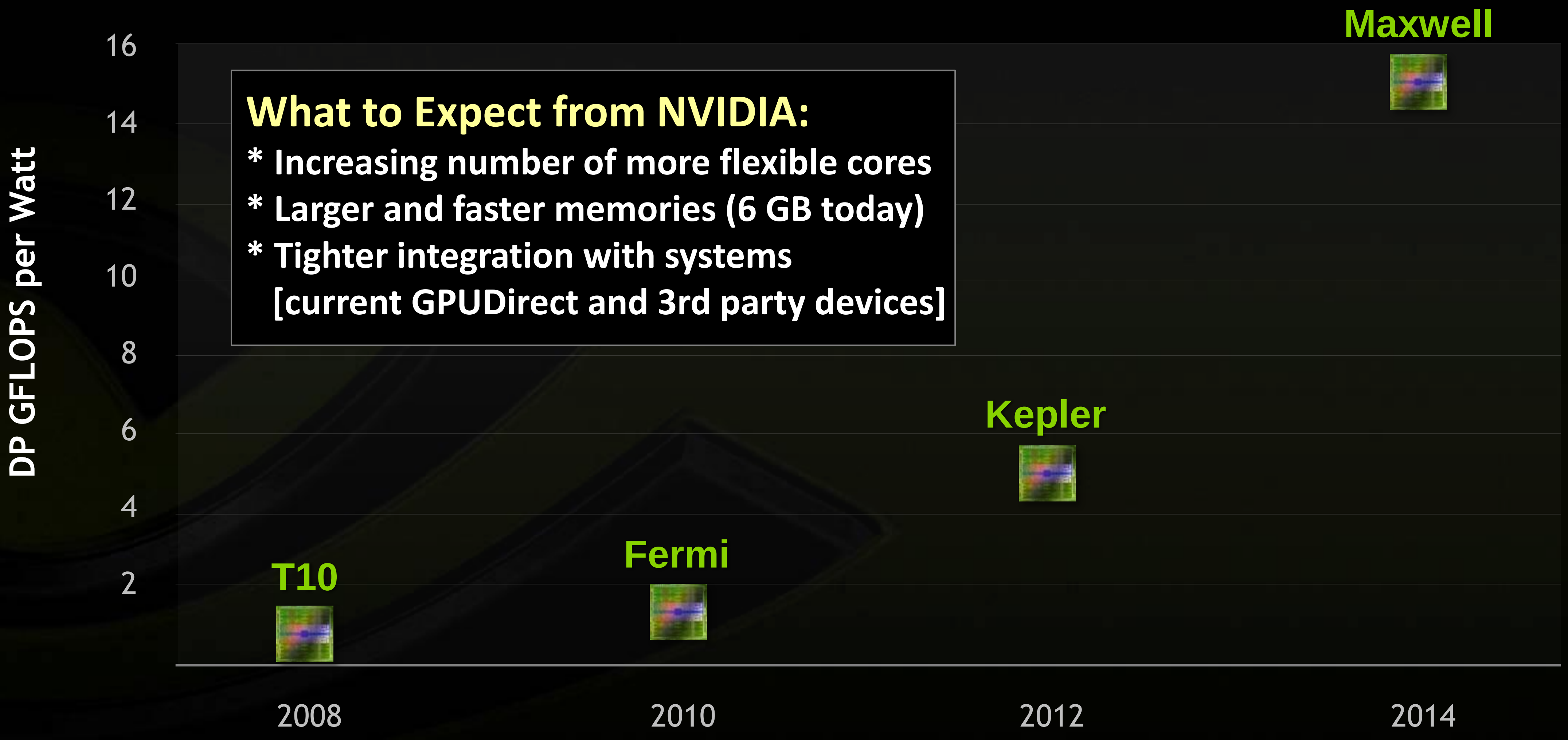
http://www.engadget.com/2011/01/05/nvidia-announces-project-denver-arm-cpu-for-the-desktop/

An ARM processor coupled with an NVIDIA GPU represents the computing platform of the future. A high-performance CPU with a standard instruction set will run the serial parts of applications and provide compatibility while a highly-parallel, highly-efficient GPU will run the parallel portions of programs.

BY BILL DALLY
Posted on Jan 5 2011 at 01:05:16 PM in Mobile
VIEW COMMENTS

“PROJECT DENVER” PROCESSOR TO USHER IN NEW ERA OF COMPUTING

NVIDIA Tesla CUDA GPU Roadmap



What to Expect from NVIDIA:

* Increasing number of more flexible cores
* Larger and faster memories (6 GB today)
* Tighter integration with systems
  [current GPUDirect and 3rd party devices]

Top System Vendors for Climate-Weather and GPUs

Dedicated Weather-Climate Systems

(TAKEN FROM THE NOVEMBER 2010 LIST OF TOP500 SUPERCOMPUTER SITES)

Worldwide Ranking	Organization	Country	Peak Teraflops	**Sustained Teraflops**	Supplier
# 19	KMA	Korea	379.01	**316.40**	CRAY XE6
# 20	KMA	Korea	379.01	**316.40**	CRAY XE6
# 32	NOAA/ORNL	USA	259.66	**194.40**	CRAY XT6
# 50	NOAA/ESRL	USA	148.12	**126.50**	Aspen Cluster
# 56	JAMSTEC	JAPAN	131.07	**122.40**	NEC SX9
# 57	ECMWF	UK	156.42	**115.90**	IBM Power 575
# 40	ECMWF	UK	156.42	**115.90**	IBM Power 575
# 58	DKRZ	GY	151.60	**115.90**	IBM Power 575
# 81	NAVO	USA	117.14	**90.84**	CRAY XT5
# 93	NAVO	USA	102.27	**78.68**	IBM Power 575
#101	NIES	JAPAN	177.12	**74.84**	HP Cluster
#103	NCEP	USA	93.85	**73.06**	IBM Power 575
#104	NCEP	USA	93.85	**73.06**	IBM Power 575
#127	NCAR	USA	76.40	**59.68**	IBM Power 575

CRAY THE SUPERCOMPUTER COMPANY

IBM

System providers now offer GPUs:

- **Cray XK6**
- **IBM iDataPlex**

Recent NVIDIA and Cray Alliance Announced

Several Cray Centers Dedicated Climate and Weather HPC

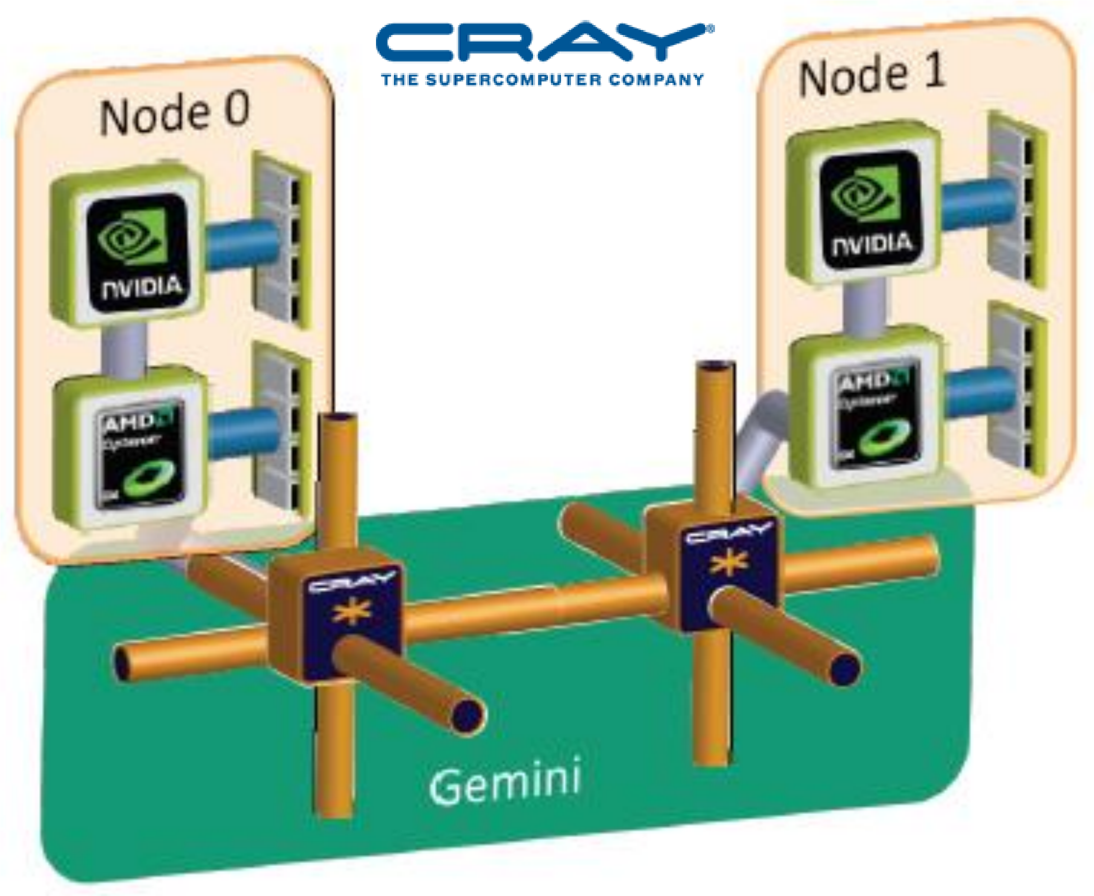
Many have GPU Evaluations Ongoing

Key Joint ORNL HPC projects that include C-W models on GPU:
- **Keeneland**
- **Titan (20 PF)**

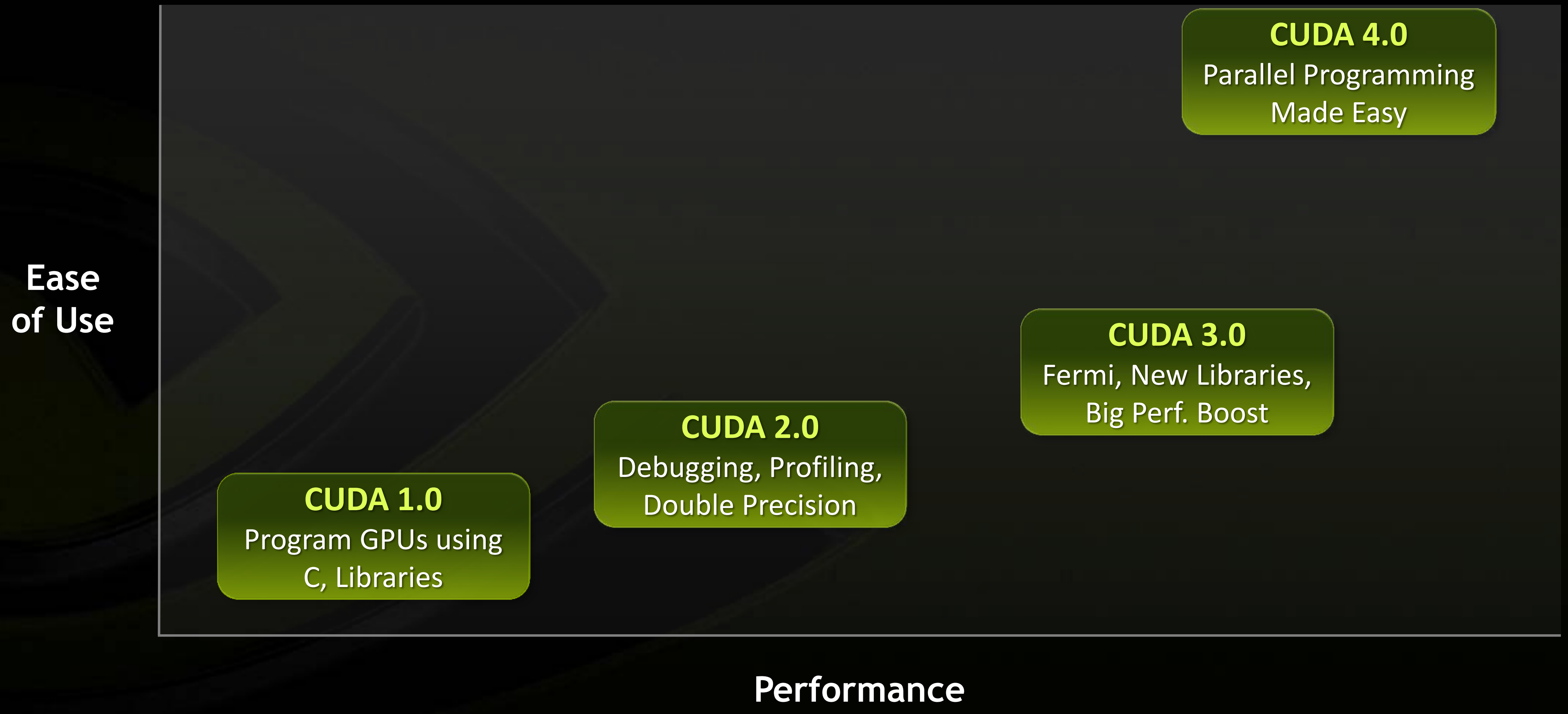
Data Courtesy of Mr. Per Nyberg, Cray Inc.

Cray Development of the XK6 with NVIDIA GPUs

XK6 Compute Node Characteristics	
Host Processor	AMD Series 6200 (Interlagos)
Host Processor	“TBA” Performance
Tesla X2090 Cores	512
Tesla X2090 Perf.	665 Gflops
Host Memory	16, 32, or 64GB 1600 MHz DDR3
Tesla X090 Memory	6GB GDDR5 capacity 170 GB/sec
Gemini High Speed Interconnect	
Upgradeable to KEPLER many-core processor	



NVIDIA CUDA Software Roadmap



NVIDIA CUDA Overview

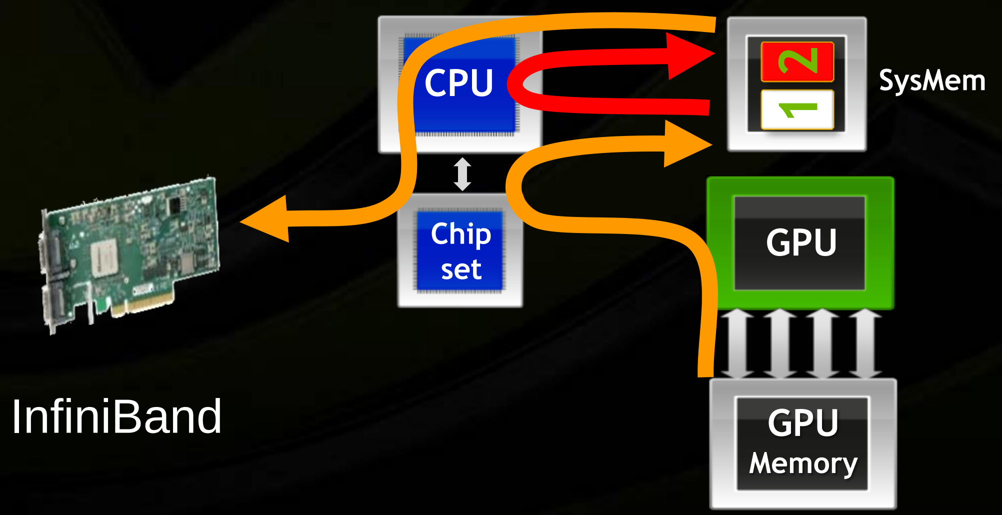
	Platform	Programming Model	Libraries	Tools
New in CUDA 4.0	GPUDirect 2.0 Fast Path to Data	Unified Virtual Addressing C++ new/delete C++ Virtual Functions	Thrust C++ Library Templated Perf Primitives	Parallel Nsight Pro
	Hardware Support ECC Memory Double Precision Native 64-bit Architecture Concurrent Kernel Execution Dual Copy Engines Multi-GPU support 6GB per GPU supported **Operating System Support** MS Windows 32/64 Linux 32/64 support Mac OSX support **Cluster Management** NVIDIA GPUDirect Tesla Compute Cluster (TCC) Graphics Interoperability	**C support** NVIDIA C Compiler CUDA C Parallel Extensions Function Pointers Recursion Atomics malloc/free **C++ support** Classes/Objects Class Inheritance Polymorphism Operator Overloading Class Templates Function Templates Virtual Base Classes Namespaces **Fortran, OpenCL**	**NVIDIA Library Support** Complete math.h Complete BLAS Library Sparse Matrix Math Library RNG Library FFT Library (1D, 2D and 3D) Image Processing Library (NPP) Video Processing Library (NPP) **3rd Party Math Libraries** CULA Tools MAGMA IMSL VSIPL	**NVIDIA Tools Support** Parallel Nsight 1.0 IDE cuda-gdb multi-GPU debugger CUDA/OpenCL Visual Profiler CUDA Memory Checker CUDA C SDK CUDA Disassembler **CUDA Partner Tools** Allinea DDT RogueWave /Totalview Vampir Tau CAPS HMPP

NVIDIA GPUDirect
Accelerated Communication with Network and Storage Devices

Without GPUDirect

Same data copied three times:

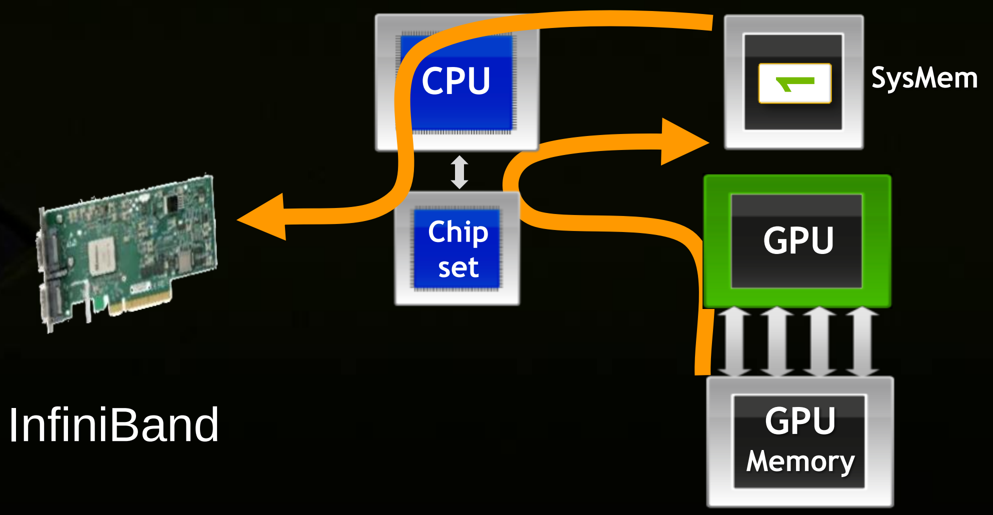
1. GPU writes to pinned sysmem1
2. CPU copies from sysmem1 to sysmem2
3. InfiniBand driver copies from sysmem2



With GPUDirect

Data only copied twice

Sharing pinned system memory makes sysmem-to-symem copy unnecessary



Summary

- Several C-W Models support NVIDIA CUDA and GPUs
  - Mostly dynamics today but full implementations coming
- Joint Collaboration with Key Organizations Ongoing
  - Collaboration with C-W research organizations and vendors
- Learn More About NVIDIA HPC Solutions
  - More at: www.nvidia.com
  - Want to investigate GPUs, please contact sposey@nvidia.com

Thank You, Questions ?

Stan Posey | HPC Industry Development

NVIDIA, Santa Clara, CA, USA